and for characterizing nucleic acid molecules by synthesizing nucleic acid molecules containing at least one non-canonical nucleotide *in vitro*.

In one preferred embodiment, the invention uses mutant RNA polymerases that efficiently utilize deoxynucleoside triphosphates as substrates. *In vitro*, this mutant will synthesize RNA, DNA, or 'transcripts' of mixed dNMP/rNMP composition from a template molecule depending on the mix of NTPs or dNTPs present in the synthesis reaction.

Mutant Polymerases of the Present Invention

In a preferred embodiment, the polymerase mutation is conservative, for example, changing tyrosine 639 (of T7 polymerase) within the active site to phenylalanine, and does not substantially affect promoter specificity or overall activity. Non-conservative mutations of this tyrosine also reduce discrimination between deoxy- and ribo-nucleoside triphosphates, but these mutations also typically cause large activity reductions. Among the most active of the non-conservative mutations, enzymes with methionine or leucine in place of the wild-type tyrosine at the 639 position of T7 RNAP had about half the enzymatic activity of the wild-type enzyme.

Of 26 other mutations of other amino acid positions examined in and around the active site of T7 RNAP, none showed marked effects on rNTP/dNTP discrimination.

T7 RNA polymerase can use RNA templates as well as DNA templates and is capable of both primer extension and *de novo* initiation. The Y639F mutant, described below in the Examples, retains the ability to use RNA or DNA templates. Thus, this mutant can display *de novo* initiated or primed DNA directed DNA polymerase, reverse transcriptase, RNA directed RNA polymerase, or DNA directed RNA polymerase

activities depending on the templates and substrates presented to it in the synthesis reaction.

A major theme of research on nucleic acid polymerases over the past several years has been the discovery of extensive structural similarity between the majority of these enzymes, even those from functionally different classes (Sousa, et al., 1993; Pelletier, et al., 1994; Jacob-Molina, et al., 1993; Kohlstaedt, et al., 1992; Sawaya, et al., 1994; Steitz, et al., 1994). One part of this work has been the identification of well-conserved residues. Five amino acids have been identified as invariant in a large number of DNA-directed RNA polymerases (Delarue, et al., 1990). In T7 RNAP these are D537, K631, Y639, G640A and D812. A specific, conserved function has been revealed for the two invariant aspartates in coordinating the catalytic Mg^{++} ion (Sousa, et al., 1993; Pelletier, et al., 1994; Jacob-Molina, et al., 1993; Kohlstaedt, et al., 1993; Sawaya, et al., 1994; Steitz, et al., 1994). Our observations imply a similarly specific and conserved function for Y639 as a sensor of inappropriate geometry or structure in the template-NTP-primer/RNA complex.

The present invention encompasses methods for synthesis of nucleic acids containing at least one non-canonical nucleotide using mutant nucleic acid polymerases which have reduced discrimination for non-canonical nucleoside triphosphate substrates. The examples below demonstrate the reduced dNTP/rNTP discrimination of mutants of T7 RNA polymerase and SP6 RNA polymerase. Genes for other polymerases may be modified or mutated to obtain mutant enzymes which have similar reduced discrimination for non-canonical substrates. If one wished to mutate an RNA

polymerase to have the properties described herein, one would first locate the amino acid corresponding to the T7 polymerase Y639 in other RNA polymerases. Identification of the corresponding mutation site in other polymerases can be done by the well-established procedure of sequence alignment, which involves aligning the amino acid sequences of two proteins, introducing gaps and insertions, and shifting the sequences with respect to each other while maintaining their original linearity. Such alignment procedures are often performed with the aid of one or more computer programs into which the amino acid sequences that one wishes to compare have been entered. When the sequence identity of two proteins is high enough (greater than or equal to 30%) over a sufficient length of amino acids (greater than or equal to 50), this procedure is very reliable in identifying amino acids that occupy corresponding structural and functional positions in the two proteins. Such conditions are met for the T7-type group of RNAPs, which include T7, T3, ϕI, ϕIIH, W31, gh1, Y, A1122, SP6 and mitochondrial RNAPs, and allow identification of the mutation site corresponding to Y639 in T7 RNAP.

Using this method, we predicted that the amino acid in w.t. SP6 RNAP that corresponded to the Y639 site in T7 RNAP was Y631, and as described herein, mutagenesis of this site resulted in a Y631F mutant SP6 RNAP which has a similar reduced discrimination for dNTPs compared to rNTPs like the Y639F mutant T7 RNAP.

From alignment studies, it is known that there is a conserved motif present in T7-like RNAPs and class I DNAPs with the following consensus sequences:

... K — — — — — — — Y G ...